Notes on Number Theory and Discrete Mathematics
Print ISSN 1310–5132, Online ISSN 2367–8275
XXXX, Volume XX, Number X, XXX–XXX
DOI: 10.7546/nntdm.XXXX

On ternary Dejean words avoiding 010

Pascal Ochem

LIRMM, Université de Montpellier, CNRS
Montpellier, France
ochem@lirmm.fr

Received: DD Month XXXX **Revised:** DD Month XXXX
Accepted: DD Month XXXX **Online First:** DD Month XXXX

Abstract: Thue has shown the existence of three types of infinite square-free words over $\{\texttt{0},\texttt{1},\texttt{2}\}$ avoiding the factor $\texttt{010}$. They respectively avoid $\{\texttt{010},\texttt{212}\}$, $\{\texttt{010},\texttt{101}\}$, and $\{\texttt{010},\texttt{020}\}$. Also Dejean constructed an infinite $\left(\frac{7}{4}^+\right)$-free ternary word. A word is d-directed if it does not contain both a factor of length d and its mirror image. We show that there exist exponentially many $\left(\frac{7}{4}^+\right)$-free 180-directed ternary words avoiding $\texttt{010}$. Moreover, there does not exist an infinite $\left(\frac{7}{4}^+\right)$-free 179-directed ternary word avoiding $\texttt{010}$.

Keywords: Word, Repetitions.
2020 Mathematics Subject Classification: 68R15.

1 Introduction

This note is about words avoiding repetitions, a well-studied area in combinatorics on words [5,6]. A *repetition* is a factor of the form $r = u^n v$ where u is non-empty and v is a prefix of u. Then $|u|$ is the *period* of the repetition r and its *exponent* is $|r|/|u|$. A word is α^+*-free* (resp. α*-free*) if it contains no repetition with exponent β such that $\beta > \alpha$ (resp. $\beta \geqslant \alpha$).

We consider ternary square-free words (i.e., 2-free words) with additional avoidance constraints. Thue [7] has shown that there exist infinite square-free words avoiding the factors $\{\texttt{010},\texttt{212}\}$,

Copyright © 2023 by the Authors. This is an Open Access paper distributed under the terms and conditions of the Creative Commons Attribution 4.0 International License (CC BY 4.0). https://creativecommons.org/licenses/by/4.0/

$\{\texttt{010}, \texttt{101}\}$, and $\{\texttt{010}, \texttt{020}\}$ respectively. In each case, the avoiding word is essentially unique. Consider for example the fixed point of the morphim $\texttt{0} \mapsto \texttt{012}$, $\texttt{1} \mapsto \texttt{02}$, $\texttt{2} \mapsto \texttt{1}$, which is a famous ternary word named b_3 in [1]. Then b_3 avoids squares and $\{\texttt{010}, \texttt{212}\}$, but for every finite factor f of b_3, there exist only finitely many ternary words avoiding squares and $\{\texttt{010}, \texttt{212}, f\}$. See also [1] for the construction of the other two morphic words. Thus the language L of (finite or infinite) ternary square-free words avoiding $\texttt{010}$ is worth considering since it contains these three interesting words.

To refine square-freness, Dejean [3] introduced the notion of *repetition threshold* $RT(n)$, which is the least real α such that there exists an infinite word over the n-letter alphabet that is α^+-free. She proved that $RT(3) = \frac{7}{4}$ by exhibiting a 19-uniform morphism whose fixed point is $\left(\frac{7}{4}^+\right)$-free. Our main result shows that both constraints, avoiding $\texttt{010}$ and being $\left(\frac{7}{4}^+\right)$-free, can be satisfied simultaneously. Moreover, the language L' of ternary $\left(\frac{7}{4}^+\right)$-free avoiding $\texttt{010}$ still contains exponentially many words. Notice that L' is also closed under reversal. However, there exist infinite words $w \in L'$ such that the factor set of w is not closed under reversal, as opposed to the three cases studied by Thue. A quantitative notion of "factor set not closed under reversal" is defined in [2] as follows. A word w is *d-directed* if for every factor f of w of length d, the reversed word f^R is not a factor of w.

Theorem 1.1.

- *There exist exponentially many $\left(\frac{7}{4}^+\right)$-free* 180*-directed ternary words avoiding* $\texttt{010}$.
- *There does not exist an infinite $\left(\frac{7}{4}^+\right)$-free* 179*-directed ternary word avoiding* $\texttt{010}$.

2 Proof

A morphism is *q-uniform* if the image of every letter has length q. Also, a q-uniform morphism $h : \Sigma_s^* \to \Sigma_e^*$ is *synchronizing* if for any $a, b, c \in \Sigma_s$ and $v, w \in \Sigma_e^*$, if $h(ab) = vh(c)w$, then either $v = \varepsilon$ and $a = c$ or $w = \varepsilon$ and $b = c$. We will need the following result, which corresponds to the case $n = 1$ of the main lemma in [4].

Lemma 2.1. *Let $\alpha, \beta \in \mathbb{Q}$, $1 < \alpha < \beta < 2$. Let $h : \Sigma_s^* \to \Sigma_e^*$ be a synchronizing q-uniform morphism (with $q \geqslant 1$). If $h(w)$ is β^+-free for every α^+-free word w such that $|w| < \max\left(\frac{2\beta}{\beta-\alpha}, \frac{2(q-1)(2\beta-1)}{q(\beta-1)}\right)$, then $h(t)$ is β^+-free for every (finite or infinite) α^+-free word t.*

Let us prove the first part of Theorem 1.1. We use lemma 2.1 to show that the image of every $\left(\frac{7}{5}^+\right)$-free word over Σ_4 by the following 1557-uniform morphism is $\left(\frac{7}{4}^+\right)$-free. With $\alpha = \frac{7}{5}$ and $\beta = \frac{7}{4}$, we get $\frac{2\beta}{\beta-\alpha} = 10$, so we need to check the $\left(\frac{7}{4}^+\right)$-freeness of the image of every

$\left(\frac{7}{5}^+\right)$-free word of length 9. It is easy to check that such words avoid `010` and are 180-directed.

```
0 →  p21020121021201210120210201202120121021202102012021012102120121020120212012101202102012102120210121020120212012102120 21
     02012021012102120121012021020120212012102120210201210120212012102012021012102120210201202120121012021020121021202101210
     20120212012102120210201210120212012102012021012102120121012021020121021202101210201202120121021202102012021012102120121
     02012021201210120210201210212021012102012021201210212021020121012021201210201202101210212012101202102012021201210212021
     02012021012102120121020120212012101202102012102120210121020120212012102120210201202101210212012101202102012021201210212
     02102012101202120121020120210121021201210120210201210212021012102012021201210212021020120210121021201210201202120121012
     02102021021201210201202120121012021020121021202101210201202120121021202102012021012102120121012021020120212012102120210
     2012101202120121020120210121021202102012021201210120210201210212021012102012021201210212021020121012021201210 2012021
1 →  p21020121021201210120210201202120121021202102012021012102120121020120212012101202102012102120210121020120212012102120 21
     02012021012102120121012021020120212012102120210201210120212012102012021012102120121012021020121021202101210201202120121
     02120210201202101210212012102012021201210120210201210212021012102012021201210212021020121012021201210201202101210212012
     10120210201202120121021202102012021012102120121020120212012101202102012102120210121020120212012102120210201202101210212
     01210120210201202120121021202102012101202120121020120210121021202102012021201210120210201210212021012102012021201210212
     02102012101202120121020120210121021201210120210201210212021012102012021201210212021020120210121021201210201202120121012
     02102012102120210121020120212012102120210201210120212012102012021012102120121012021020120212012102120210201202120211020
     12021012102120121020120212012101202102012102120210121020120212012102120210201202101210212012101202102012021201210212021
     020121012021201210201202101210212021020120212012101202102012102120210121020120212012102120210201210120212012102012021
2 →  p12101202102012021201210212021020120210121021201210201202120121012021020121021202101210201202120121021202102012021012 10
     21201210120210201202120121021202102012101202120121020120210121021202102012021201210120210201210212021012102012021201210
     21202102012101202120121020120210121021201210120210201210212021012102012021201210212021020120210121021201210201202120121
     01202102012102120210121020120212012102120210201210120212012102012021012102120210201210212012101202102012021201210212021
     02012021012102120121020120212012101202102012102120210121020120212012102120210201202101210212012101202102012021201210212
     02102012101202120121020120210121021201210120210201210212021012102012021201210212021020120210121021201210201202120121012
     02102012102120210121020120212012102120210201210120212012102012021012102120121012021020120212012102120210201202101210212
     01210201202120121012021020121021202101210201202120121021202102012021012102120121012021020120212012102120210201210120212
     0121020120210121021202102012021201210120210201210212021012102012021201210212021020121012021201210201202101210212 02102
3 →  p12101202102012021201210212021020120210121021201210201202120121012021020121021202101210201202120121021202102012021012 10
     21201210120210201202120121021202102012101202120121020120210121021201210120210201210212021012102012021201210212021020120
     21012102120121020120212012101202102012102120210121020120212012102120210201210120212012102012021012102120210201210212012
     10120210201202120121021202102012021012102120121020120212012101202102012102120210121020120212012102120210201202101210212
     01210120210201202120121021202102012101202120121020120210121021202102012021201210120210201210212021012102012021201210212
     02102012101202120121020120210121021201210120210201210212021012102012021201210212021020120210121021201210201202120121012
     02102012102120210121020120212012102120210201210120212012102012021012102120121012021020120212012102120210201202101210212
     01210201202120121012021020121021202101210201202120121021202102012021012102120121012021020120212012102120210201210120212
     0121020120210121021202102012021201210120210201210212021012102012021201210212021020121012021201210201202101210212 02102
```

with the common prefix

```
p =  01210212012101202102012021201210212021020120210121021201210201202120121012021020121021202101210201202120121021202102012
     02101210212012101202102012021201210212021020121012021201210201202101210212012101202102012102120210121020120212012102120
     21020121012021201210201202101210212021020121021201210120210201202120121021202102012021012102120121012021020121021202101
     21020120212012102120210201202101210212012102012021201210120210201210212021012102012021201210212021020121012021201210201
     2021012102120
```

The C code to find and check this morphism is available at https://www.lirmm.fr/~ochem/morphisms/dejean010.htm.

Now we prove the second part of Theorem 1.1. Let L_{179} be the language of $\left(\frac{7}{4}^+\right)$-free 179-directed ternary words avoiding `010`. The standard backtracking algorithm is rather slow to show that L_{179} is finite. So we use the following simple optimisation. We choose an arbitrary small factor in L_{179}, namely f = `10121021201210201202120121012 02`. Then we use backtracking to show that:

(1) The language of words in L_{179} avoiding f is finite.

(2) The language of words in L_{179} with prefix f is finite.

This rules out that L_{179} contains an infinite word w: by item (1), w must contain f, but then by item (2), w cannot be extended indefinitely to the right of this factor f.

References

[1] G. Badkobeh and P. Ochem. Characterization of some binary words with few squares. *Theoret. Comput. Sci.* **588** (2015), 73–80.

[2] G. Badkobeh, T. Harju, P. Ochem, and M. Rosenfeld. Avoiding square-free words on free groups. *Theoretical Computer Science* **922** (2022), 206–217.

[3] F. Dejean. Sur un théorème de Thue, *J. Combin. Theory. Ser. A*, 13:90–99, 1972.

[4] P. Ochem. A generator of morphisms for infinite words. *RAIRO - Theoret. Informatics Appl.*, 40:427–441, 2006.

[5] P. Ochem, M. Rao, and M. Rosenfeld. Avoiding or limiting regularities in words. *Sequences, Groups, and Number Theory (Trends in Mathematics)*, 177–212, 2018.

[6] N. Rampersad and J. Shallit. Repetitions in words. *Combinatorics, Words and Symbolic Dynamics (Encyclopedia of Mathematics and its Applications)*, 101–150, 2012.

[7] A. Thue. Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen. *Norske vid. Selsk. Skr. Mat. Nat. Kl.* **1** (1912), 1–67. Reprinted in *Selected Mathematical Papers of Axel Thue*, T. Nagell, editor, Universitetsforlaget, Oslo, (1977), 413–478.